Application d'une nouvelle classe de lagrangiens augmentés en contrôle optimal de systèmes distribués

J. Frederic Bonnans

IRIA

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél: 954 90 20

Rapports de Recherche

N° 102

APPLICATION D'UNE NOUVELLE CLASSE DE LAGRANGIENS AUGMENTÉS EN CONTRÔLE OPTIMAL DE SYSTÈMES DISTRIBUÉS

Joseph Fréderic BONNANS

Novembre 1981

APPLICATION D'UNE NOUVELLE CLASSE DE LAGRANGIENS

AUGMENTES EN CONTROLE OPTIMAL DE SYSTEMES DISTRIBUES

Joseph Frédéric BONNANS

RESUME

Ce rapport présente une nouvelle méthode de résolution des problèmes de contrôle optimal de systèmes distribués. On suppose le critère convexe et l'équation d'état affine par rapport à la paire (état, contrôle). L'équivalence entre le problème initial et la minimisation par rapport au contrôle, à l'état et aussi à l'état adjoint d'un lagrangien augmenté est démontrée. Le lagrangien augmenté est la somme du lagrangien et de deux termes pénalisant l'équation d'état et l'équation de l'état adjoint. La méthode est appliquée à un exemple concret et mise en oeuvre numériquement.

ABSTRACT

This paper describes a new method to solve optimal control problems of distributed parameter systems. The study is restricted to problems with a convex criterion and a state equation affine with respect to the pair (state, control). We show the equivalence between the initial problem and the minimisation with respect to the control, the state and also the costate of some augmented lagrangian, obtained by addition of the lagrangian and of two terms penalizing the state and costate equations. An example is treated, and numerical results are given.

INRIA, Domaine de Voluceau, BP 105, Rocquencourt, 78153 LE CHESNAY CEDEX

PLAN

I. INTRODUCTION

Un problème de contrôle est l'étude de la minimisation d'un critère par rapport à une variable appelée contrôle, le critère dépendant d'une fonction du contrôle appelée état. Une équation d'état permet, pour un contrôle donné, de calculer l'état. Lorsque le système est distribué, l'état et quelquefois le contrôle appartiennent à des espaces vectoriels de dimension infinie.

La méthode classique pour résoudre un tel problème consiste à introduire un état adjoint qui permet le calcul du gradient du critère par rapport au contrôle. Il est alors possible de résoudre le problème par une méthode de descente.

Ceci s'interprète comme une méthode de gradient réduit (D. Gabay [9]) si on considère le couple état-contrôle comme la variable inconnue, soumise à une contrainte : l'équation d'état. Dans ce cadre général se situe l'étude de J.P. Yvon [16] sur l'application de la méthode de pénalisation au contrôle. Il est apparu que le coefficient de pénalisation nécessaire pour assurer une bonne approximation de la solution exacte est assez important et peut donc causer un mauvais conditionnement de la fonctionnelle pénalisée.

Dans le même cadre il est naturel de penser à appliquer des méthodes de lagrangien augmenté. Cette théorie, issue des travaux de M.R. Hestenes [11] et M.J.D. Powell [13] a déjà été appliquée à des problèmes convexes en dimension infinie : cf. R. Glowinski [10] et M. Fortin, R. Glowinski, F. Thomasset [8]. G. Di Pillo et L. Grippo [3] ont traité le cas d'un problème de contrôle en recherchant par l'algorithme d'Uzawa les points-selle du lagrangien augmenté. Leurs résultats ne permettent pas de conclure sur l'efficacité pratique de la méthode dans la mesure où ils ne précisent pas les temps de calcul.

Récemment ont été introduits d'autres types de lagrangiens augmentés. Décrivons les dans le cadre d'un problème général d'optimisation en dimension finie. Soit n et m deux entiers positifs tels que $n > m$, f une application de R^n vers R et g une application de R^n vers R^m. Soit le problème :

$$\begin{cases} \text{minimiser } f(x), x \in R^n, \\ \text{sous la contrainte } g(x) = 0. \end{cases}$$

Le lagrangien associé à ce problème est

$$\ell(x,\lambda) = f(x) + (\lambda,\ g(x)),$$

λ étant la variable duale. Les conditions nécessaires d'optimalité pour que $(\bar{x},\bar{\lambda})$ soit point-selle du lagrangien sont :

$$\frac{\partial \ell}{\partial \lambda}(\bar{x},\bar{\lambda}) = g(\bar{x}) = 0,$$

$$\frac{\partial \ell}{\partial x}(\bar{x},\bar{\lambda}) = \frac{\partial f}{\partial x}(\bar{x},\bar{\lambda}) + \frac{{}^t\partial g}{\partial x}(\bar{x}).\ \bar{\lambda} = 0.$$

Habituellement on obtient le lagrangien augmenté en pénalisant la première condition d'optimalité. Un prolongement possible consiste à traiter les deux conditions d'optimalité de façon symétrique ce qui donne :

$$L(x,\lambda) = f(x) + (\lambda,\ g(x)) + c||g(x)||^2 + d||\frac{\partial f}{\partial x}(x) + \frac{{}^t\partial g}{\partial x}(x).\lambda||^2.$$

P.T. Boggs et J.W. Tolle [2] ont étudié le cas où $c > 0$ et $d < 0$. Ils remarquent qu'un point-selle de ℓ vérifiant la condition suffisante d'optimalité est encore, si d est assez petit, point-selle de L. Posons

$$M(x) = \sup_{\lambda} L(x,\lambda)$$

alors si $\frac{{}^t\partial g}{\partial x}$ est partout de rang plein, M(x) est partout fini et

$$M(x) = L(x,\ \lambda(x))$$

où λ est solution d'un système linéaire. Le problème de départ équivaut à la minimisation de la fonction sans contraintes M, qui apparaît donc comme une fonction de pénalisation exacte.

G. Di Pillo et L. Grippo [4] ont étudié le cas où $d > 0$. Ils montrent, sous les hypothèses classiques en optimisation, que si c est assez grand, un point-selle de ℓ sera minimum local strict de L. On est donc ramené au problème de la minimisation de L par rapport à x <u>et</u> λ. G. Di Pillo, L. Grippo et F. Lampariello [5] ont appliqué cette méthode à un problème de contrôle optimal. Ils minimisent L par une méthode de Newton en profitant de la structure creuse du Hessian de L.

Cet article étend certains des résultats de G. Di Pillo et L. Grippo [4] dans le cadre de problèmes de contrôle en dimension infinie. Après avoir précisé le problème et les hypothèses - en particulier la convexité du critère et la linéarité de l'équation d'état - on prouve l'identité entre les points-selle du lagrangien (qui correspondent à une solution du problème de contrôle) et l'argument du minimum du lagrangien augmenté. L'exemple du contrôle d'un système régi par l'équation de la chaleur est traité pour une observation distribuée ou finale. Des résultats numériques complètent l'étude et donnent des indications sur l'importance du choix de la norme de l'état, et sur le cas où l'équation d'état n'est pas linéaire.

II. RESULTATS THEORIQUES ET APPLICATIONS

II.1. Définition du problème

Le problème de contrôle est présenté ici dans un cadre abstrait. Ceci permet d'inclure les cas où l'équation d'état est de type elliptique, parabolique ou hyperbolique avec une observation distribuée ou finale.

Tous les espaces de fonction sont des espaces de Hilbert. Notons E' le dual d'un espace E. Soit le contrôle $u \in U$ et l'état $y \in Y$. L'équation d'état est

(1) $\quad Ay = f + Bu$ dans W,

espace de l'équation d'état, où $f \in W$ est donné, $B \in \mathcal{L}(U,W)$ et A est un isomorphisme de Y sur W ; c'est-à-dire que A^{-1} existe, et que A et A^{-1} sont continus.

Remarquons que l'équation (1) est bien posée : chaque contrôle u définit un état y unique, l'état dépendant continûment du contrôle. Le problème de contrôle optimal est :

$$(2)\qquad \begin{cases} \text{minimiser } J(y,u) = F(y) + G(u) \\ \\ \text{pour } (y,u) \in Y \times U \text{ vérifiant (1).} \end{cases}$$

F et G sont des fonctions convexes semi-continues inférieurement de Y dans R et de U dans R. On suppose que le problème de contrôle (2) a une solution : il existe $(\bar{y},\bar{u}) \in Y \times U$ tels que :

$$(3)\qquad \begin{cases} J(\bar{y},\bar{u}) \in R \\ \\ J(\bar{y},\bar{u}) \leq J(y,u) \quad \forall (y,u) \in Y \times U \text{ vérifiant (1)} \\ \\ \text{et } (\bar{y},\bar{u}) \text{ vérifie (1).} \end{cases}$$

De plus on fait l'hypothèse sur F :

$$(4)\qquad \begin{cases} F \text{ est Gâteaux-différentiable sur } Y \text{ et il existe } \alpha > 0 \text{ tel que} \\ \\ \text{pour tout } y_1,\, y_2 \in V \text{ on ait :} \\ \\ F(y_2) \geq F(y_1) + \langle F'(y_1),\, y_2 - y_1 \rangle_{Y'Y} + \alpha \| F'(y_1) - F'(y_2) \|^2_{Y'}\ . \end{cases}$$

Dans (4) $F'(y)$ est la G-différentielle de F au point y.

<u>REMARQUE 1</u> (4) n'implique pas que F soit coercive ou strictement convexe : en effet une fonction F affine vérifie (4) ; c'est également le cas si

$$F(y) = \frac{1}{2} \| Cy - z_d \|^2_X ,$$

où X est un Hilbert, $C \in \mathcal{L}(Y, X)$ et $z_d \in X$. On a alors :

$$F(y_2) = F(y_1) + \langle F'(y_1), y_2 - y_1\rangle_{Y'Y} + \frac{1}{2}||C(y_2 - y_1)||^2_X$$

D'autre part,

$$F'(y_2) - F'(y_1) = C^* C(y_2 - y_1).$$

C étant continue, il existe m > 0 tel que

$$||F'(y_2) - F'(y_1)||^2_{Y'} \leq m||C(y_2 - y_1)||^2_X.$$

donc (4) est satisfaite pour $\alpha \leq 1/2\ m$.

Notons que si F_1 et F_2 vérifient (4), $F_1 + F_2$ vérifie aussi (4). On verra, sur un exemple, que la pénalisation de la partie positive de y vérifie aussi (4).

Au problème (2) est associé le lagrangien $L : Y \times U \times W' \to R$ défini par

$$L(y,u,p) = F(y) + G(u) - \langle p, Ay - f - Bu\rangle_{W'W}.$$

Le signe - est introduit pour retrouver ultérieurement l'état adjoint tel qu'il est défini dans le livre de J.L. Lions [12]. On note A^*, B^* les opérateurs adjoints à A et B et $\partial G(u)$ le sous-différentiel de G au point u :

$$\partial G(u) = \{u^* \in U' ;\ G(v) \geq G(u) + \langle u^*, v - u\rangle_{U'U} \quad \forall v \in U\}.$$

De l'étude générale des points-selle (I. Ekeland, R. Temam [6] p. 158) on déduit la

<u>PROPOSITION 1</u> $(\bar{y},\bar{u}) \in (Y \times U)$ est solution du problème (2) si et seulement si il existe $\bar{p} \in W'$ tel que $(\bar{y},\bar{u},\bar{p})$ est point-selle de L, et alors on a :

(5) $\quad \partial G(\bar{u}) \ni - B^*\bar{p},$

(6) $\quad A\bar{y} - f - B\bar{u} = 0,$

(7) $\quad A^*\bar{p} - F'(\bar{y}) = 0.$

II.2. L'approche par le lagrangien augmenté

Soit le lagrangien augmenté $S : Y \times U \times W' \to R$ défini par

$$S(y,u,p) = L(y,u,p) + c_1||Ay - f - Bu||^2_W + c_2||A^*p - F'(y)||^2_{Y'}.$$

S s'obtient en ajoutant au lagrangien deux termes pénalisant l'équation d'état (6) et l'équation de l'état adjoint (7). Voici l'énoncé du principal résultat :

THEOREME : sous les hypothèses (3) et (4), étant donné $c_2 > 0$, il existe $c^* > 0$ tel que, si $c_1 > c^*$, l'ensemble des (y,u,p) minimisant S est identique à l'ensemble des points-selle de L.

PREUVE. Soit (y,u,p) un point-selle de L et posons

$$\Delta = S(\bar{y} + \delta y, \bar{u} + \delta u, \bar{p} + \delta p) - S(\bar{y},\bar{u},\bar{p}).$$

On désire montrer que $\Delta \geq 0$, Δ n'étant nul que si $(\bar{y} + \delta y, \bar{u} + \delta u, \bar{p} + \delta p)$ est aussi point-selle de L. D'après (6) et (7) on a

$$S(\bar{y},\bar{u},\bar{p}) = F(\bar{y}) + G(\bar{u}),$$

$$A(\bar{y} + \delta y) - f - B(\bar{u} + \delta u) = A\delta y - B\delta u,$$

de sorte que Δ se réduit à

$$(8) \quad \begin{cases} \Delta = F(\bar{y} + \delta y) - F(\bar{y}) + G(\bar{u} + \delta u) - G(\bar{u}) - \langle \bar{p} + \delta p, A\delta y - B\delta u\rangle_{W'W} + \\ + c_1||A\delta y - B\delta u||^2_W + c_2||A^*(\bar{p} + \delta p) - F'(\bar{y} + \delta y)||^2_{Y'}. \end{cases}$$

d'après (4), (5) et (7), on a

$$F(\bar{y} + \delta y) - F(\bar{y}) - \langle \bar{p}, A\delta y\rangle_{W'W} \geq \alpha \, ||F'(\bar{y} + \delta y) - F'(\bar{y})||^2_{Y'},$$

$$G(\bar{u} + \delta u) - G(\bar{u}) + \langle \bar{p}, B\delta u\rangle_{W'W} \geq 0.$$

En posant

$$(9) \qquad D(\delta y) = F'(\bar{y} + \delta y) - F'(\bar{y}),$$

il vient avec (7)

$$A^*(\bar{p} + \delta p) - F'(\bar{y} + \delta y) = A^*\delta p - D(\delta y) \ ;$$

d'où

$$\Delta \geq \alpha||D(\delta y)||^2_{Y'} - \langle\delta p, A\delta y - B\delta u\rangle_{W'W} + c_1||A\delta y - B\delta u||^2_W +$$

$$+ c_2||A^*\delta p - D(\delta y)||^2_{Y'}.$$

d'autre part on a $\forall\beta > 0$

$$- \langle p, A\delta y - B\delta u\rangle_{W'W} \geq - \frac{\beta}{2}||\delta p||^2_{W'} - \frac{1}{2\beta}||A\delta y - B\delta u||^2_W,$$

d'où

$$(10) \qquad \begin{cases} \Delta \geq \alpha||D(\delta y)||^2_{Y'} - \frac{\beta}{2}||\delta p||^2_{W'} + (c_1 - \frac{1}{2\beta})||A\delta y - B\delta u||^2_W + \\ + c_2||A^*\delta p - D(\delta y)||^2_{Y'}. \end{cases}$$

Posons

$$q = A^*\delta p - D(\delta y).$$

A est un isomorphisme donc A^* aussi et on a :

$$\delta p = A^{*-1}(q + D(\delta y)).$$

La continuité de A^{*-1} implique qu'il existe m > 0 tel que

$$||\delta p||^2_{W'} \leq m||q + D(\delta y)||^2_{Y'} \ ;$$

donc pour tout $\beta > 0$

$$-\frac{\beta}{2}||\delta p||^2_W \geq -m\frac{\beta}{2}||q + D(\delta y)||^2_{Y'},$$

$$\geq -m\beta\Big(||q||^2_{Y'} + ||D(\delta y)||^2_{Y'}\Big).$$

De (10) on déduit alors que

$$(11)\qquad \begin{cases} \Delta \geq (\alpha - m\beta)||D(\delta y)||^2_{Y'} + (c_1 - \frac{1}{2\beta})||A\delta y - B\delta u||^2_W + \\ + (c_2 - m\beta)||A^*\delta p - D(\delta y)||^2_{Y'}. \end{cases}$$

On choisit

$$0 < \beta < \inf\left(\frac{\alpha}{m}, \frac{c_2}{m}\right),$$

puis

$$c_1 > \frac{1}{2\beta},$$

alors les coefficients de (11) sont positifs. Supposons que $\Delta = 0$. Avec (6), (7) et la définition de $D(\delta y)$ (9) il vient :

$$(12)\qquad \begin{cases} A(\bar{y} + \delta y) - f - B(\bar{u} + \delta u) = 0 \\ A^*(\bar{p} + \delta p) - F'(\bar{y} + \delta y) = 0 \ ; \end{cases}$$

ce qui signifie que les équations de l'état et de l'état adjoint sont vérifiées. Revenant à (8) il vient dans ce cas

$$\Delta = F(\bar{y} + \delta y) - F(\bar{y}) + G(\bar{y} + \delta u) - G(\bar{u}).$$

Donc si $\Delta = 0$ on a :

$$J(\bar{y} + \delta y, \bar{u} + \delta u) = J(\bar{y}, \bar{u}) \ ;$$

ce qui signifie, avec (12), que $(\bar{y} + \delta y, \bar{u} + \delta u)$ est solution du problème de contrôle et que $\bar{p} + \delta p$ est l'état adjoint ; d'après la proposition 1 $(\bar{y} + \delta y, \bar{u} + \delta u, \bar{p} + \delta p)$ est donc point-selle de L.

REMARQUE 2. d'après la démonstration précédente, il suffit que l'hypothèse (4) soit vérifiée pour $y_1 = \bar{y}$.

D'autre part supposons qu'au lieu de (4) on ait : $\exists \alpha > 0$ tel que $\forall y_1, y_2 \in Y$

$$F(y_2) \geq F(y_1) + \langle F'(y_1), y_2 - y_1 \rangle_{Y'Y} + \alpha \| F'(y_1) - F'(y_2) \|_{Y'}^{r},$$

avec $r > 1$. Le théorème reste valable en considérant le lagrangien augmenté

$$S(y,u,p) = L(y,u,p) + c_1 \| Ay - f - Bu \|_{W}^{r'} + c_2 \| A^{*}p - F'(y) \|_{Y'}^{r},$$

r' étant défini par

$$\frac{1}{r} + \frac{1}{r'} = 1 \ ;$$

en particulier on utilisera pour démontrer ce résultat l'inégalité suivante :

$$- \langle \delta p, A\delta y - B\delta u \rangle_{W'W} \geq - \frac{\beta}{r} \| \delta p \|_{W'}^{r} - \frac{1}{\beta r'} \| A\delta y - B\delta u \|_{W}^{r'} .$$

REMARQUE 3. D'après le théorème 3 il n'est pas nécessaire d'inclure dans S un terme pénalisant la condition d'optimalité

(13) $\quad \partial G(\bar{u}) \ni - B^{*}\bar{p}.$

G. Di Pillo, L. Grippo et F. Lampariello [5] ont déjà fait cette remarque dans le cas de contrôle de systèmes de dimension finie. Bien sûr il est toujours possible de rajouter à S ce terme. Néanmoins lors des essais numériques, il est apparu que le conditionnement de S était meilleur en l'absence de terme pénalisant (13).

II.3. Lien avec la pénalisation exacte

La théorie précédente permet d'exhiber facilement une fonctionnelle pénalisée exacte. Nous revenons au problème général (2). Soit p(y) défini par

$$A^*p(y) = F'(y) ;$$

cette relation étant vérifiée par $\bar{p}$ on peut se l'imposer comme contrainte en minimisant S. Posons

$$R(y,u) = S(y,u,p(y)) ;$$

alors du théorème on déduit facilement :

PROPOSITION 2. Si c_1 est assez grand, l'ensemble des solutions du problème de contrôle (2) coïncide avec l'ensemble des (y,u) minimisant :

$$R(y,u) = J(y,u) - \langle p(y), Ay - f - Bu\rangle_{W'W} + c_1 \|Ay - f - Bu\|^2_W,$$

où $y \in Y$ et $u \in U$.

Remarquons que R est le lagrangien augmenté de M.R. Hestenes [11] dans lequel on impose à la variable duale de vérifier une condition d'optimalité. R est différentiable dans la mesure où J l'est. R. Fletcher [7] a énoncé, pour des problèmes d'optimisation en dimension finie, des résultats correspondant à la proposition 2.

II.4. Exemples d'application

Dans ce paragraphe on va appliquer l'étude précédente à un cas où le système est gouverné par une équation parabolique linéaire. On supposera d'abord l'observation distribuée puis on examinera le cas d'une observation finale.

Soit Ω un ouvert régulier de $\mathbb{R}^N$, de frontière Γ. On note $Q = \Omega \times]0,T[$

et $\Sigma = \Gamma \times]0,T[$. $H^1(\Omega)$ est défini de la façon habituelle. $L^2(\Omega)$ est identifié à son dual . Soit :

$$Z(0,T) = \{y \in L^2(0,T,H^1(\Omega)) \; ; \; \frac{dy}{dt} \in L^2(0,T,H^1(\Omega)')\},$$

muni de la norme

$$||y||_Z = (||y||^2_{L^2(0,T,H^1(\Omega))} + ||\frac{dy}{dt}||^2_{L^2(0,T,H^1(\Omega)')})^{1/2} \quad ,$$

et soit l'équation d'état : chercher $y \in Z(0,T)$ tel que :

$$(14) \qquad \begin{cases} \frac{\partial y}{\partial t} - \Delta y = f \text{ dans } Q, \\ \frac{\partial y}{\partial n} = u \text{ sur } \Sigma, \\ y(0,x) = 0 \text{ dans } \Omega. \end{cases}$$

Dans (14) $f \in L^2(Q)$ et $u \in L^2(\Sigma)$. Soit K l'ensemble des u vérifiant les contraintes sur le contrôle ; on suppose que K est convexe et fermé dans $L^2(\Sigma)$. Soit I_K sa fonction indicatrice

$$I_K(u) = \begin{cases} 0 \text{ si } u \in K, \\ +\infty \text{ sinon.} \end{cases}$$

Le problème de contrôle est, y_d étant donnée dans $L^2(Q)$:

$$(15) \qquad \begin{cases} \text{minimiser le critère} \\ J(y,u) = \frac{1}{2}||y - y_d||^2_{L^2(Q)} + \frac{N}{2}||u||^2_{L^2(\Sigma)} + I_K(u), \\ \text{pour } (y,u) \text{ vérifiant } (14). \end{cases}$$

Pour assurer l'existence d'une solution au problème (15) on prendra $N \geq 0$ si K est borné et $N > 0$ sinon.

Nous allons mettre l'équation d'état sous la forme générale (1) puis vérifier les hypothèses sur le critère, pour rentrer dans le cadre général. Puis on s'imposera certaines contraintes sur la variable duale p pour expliciter A^* et donner au lagrangien augmenté une forme numériquement utilisable.

Soit l'opérateur a : $H^1(\Omega) \to H^1(\Omega)'$ défini par

$$<a\phi,\psi>_{H^1(\Omega)'H_1(\Omega)} = \int_\Omega \nabla\phi\nabla\psi dx \quad \forall\phi,\psi \in H^1(\Omega).$$

Soit $V = \{y \in Z(0,T)\ ;\ y(0,x)\}$. On sait que $Z(0,T) \subset \mathcal{C}(0,T,L^2(\Omega))$; donc V muni de la norme de $Z(0,T)$ est un espace de Hilbert.

Soit $W = L^2(0,T,H^1(\Omega)')$ et $A : V \to W$ défini par

$$Ay = \frac{\partial y}{\partial t} + ay \text{ dans } W,$$

et soit $U = L^2(\Sigma)$ et $B \in \mathcal{L}(U,W)$ défini par

$$<Bu,\phi>_{WW'} = \int_\Sigma u\phi d\gamma.$$

Notons que $W' = L^2(0,T,H^1(\Omega))$. Alors on sait que l'équation

$$Ay = f + Bu \text{ dans } W,$$

s'interprête formellement comme (14), et que A est un isomorphisme de V dans W. Vérifions les hypothèses sur le critère. On a

$$F(y) = \frac{1}{2}||y - y_d||^2_{L^2(Q)}.$$

L'injection canonique de V dans $\mathbb{L}^2(Q)$ étant continue on peut appliquer la remarque 1 à F, qui vérifie donc (4). D'autre part

(16) $$G(u) = \frac{N}{2}||u||^2_{L^2(\Sigma)} + I_K(u).$$

K étant un convexe fermé et N positif, G est convexe s.c.i. L'existence d'un contrôle optimal dans ce cas est un résultat classique.

On rentre donc dans le cadre général et la solution du problème de contrôle s'obtient en minimisant le lagrangien augmenté

$$S(y,u,p) = \frac{1}{2}||y - y_d||^2_{L^2(Q)} + \frac{N}{2}||u||^2_{L^2(\Sigma)} - \langle p, Ay - f - Bu\rangle_{W'W}$$

$$+ c_1||Ay - f - Bu||^2_W + c_2||A^*p - (y - y_d)||^2_{Y'},$$

sous la contrainte $u \in K$, pour $(y,u,p) \in Y \times L^2(\Sigma) \times L^2(0,T,H^1(\Omega))$.

Or on sait que l'état adjoint $\bar{p}$ d'un tel système est un élément de $Z(0,T)$ tel que $\bar{p}(T,x) = 0$. On peut donc minimiser S en s'imposant les contraintes $p \in Z(0,T)$ et $p(T,x) = 0$. On a alors, pour $z \in V$:

$$\langle A^*p - (y - y_d), z\rangle_{V'V} =$$

$$= \langle -\frac{\partial p}{\partial t} + ap - (y - y_d), z\rangle_{L^2(0,T,H^1(\Omega)')L^2(0,T,H^1(\Omega))},$$

donc

$$|\langle A^*p - (y - y_d), z\rangle_{V'V}| \leq ||-\frac{\partial p}{\partial t} + ap - (y - y_d)||_{L^2(0,T,H^1(\Omega)')} \times$$

$$\times ||z||_{L^2(0,T,H^1(\Omega))}.$$

L'injection canonique de V dans $L^2(0,T,H^1(\Omega))$ étant continue $\exists\gamma > 0$ tel que

$$|\langle A^*p - (y - y_d), z\rangle_{V'V}| \leq \gamma\, ||-\frac{\partial p}{\partial t} + ap - (y - y_d)||_{L^2(0,T,H^1(\Omega)')} \times$$

$$\times ||z||_V ,$$

et ceci $\forall z \in V$ donc

$$||A^*p - (y - y_d)||_{V'} \leq \gamma||- \frac{\partial p}{\partial t} + ap - (y - y_d)||_{L^2(0,T,H^1(\Omega)')} ,$$

et on peut donc pénaliser l'équation de l'état adjoint avec la norme de $L^2(0,T,H^1(\Omega)')$ au lieu de la norme de V'. D'où :

PROPOSITION 3. Les solutions du problème de contrôle (15) s'obtiennent, si $c_2 > 0$ et c_1 est assez grand, en minimisant le lagrangien augmenté

$$\begin{aligned} S(y,u,p) &= \frac{1}{2}||y - y_d||^2_{L^2(Q)} + \frac{N}{2}||u||^2_{L^2(\Sigma)} - \\ &- \int_0^T \langle p, \frac{\partial y}{\partial t} + ay - f - Bu\rangle_{H^1(\Omega)H^1(\Omega)'} + \\ &+ c_1 \int_0^T ||\frac{\partial y}{\partial t} + ay - f - Bu||^2_{H^1(\Omega)'} + c_2 \int_0^T ||- \frac{\partial p}{\partial t} + ap - (y - y_d)||^2_{H^1(\Omega)'} \end{aligned}$$

pour $(y,u,p) \in Z(0,T) \times L^2(\Sigma) \times Z(0,T)$ et sous les contraintes :

$$\begin{cases} y(0,x) = 0 \text{ dans } \Omega, \\ u \in K, \\ p(T,x) = 0 \text{ dans } \Omega. \end{cases}$$

REMARQUE 4. On a pris une condition initiale homogène pour simplifier l'exposé. Si la condition initiale est :

(17) $$y(0,x) = y_o(x) \quad \text{dans } L^2(\Omega),$$

on se ramène au cas précédent en considérant $y_1 \in Z(0,T)$ solution de

$$\begin{cases} y_1(0,x) = y_o(x) \text{ dans } \Omega, \\ \dfrac{\partial y_1}{\partial t} + ay_1 = 0 \text{ sur } Q, \\ \dfrac{\partial y_1}{\partial n} = 0 \text{ sur } \Sigma ; \end{cases}$$

puis on prend pour nouvelle inconnue

$$z = y - y_1,$$

ce qui permet d'appliquer la proposition 3 en changeant y en z et y_d en $y_d - y_1$. Puis on revient à l'inconnue y. On voit alors que la proposition 3 reste valable dans le cas non homogène, en remplaçant la condition initiale sur y par (17).

Envisageons maintenant un cas où l'observation est la partie positive de l'état final. Soit le problème de contrôle :

$$(18)\quad \begin{cases} \text{minimiser le critère} \\ J(y,u) = \frac{1}{2}\int_\Omega [y(T,x)^+]^2 dx + N||u||^2_{L^2(\Sigma)} + I_K(u), \\ \text{pour } (y,u) \text{ vérifiant (14).} \end{cases}$$

Soit

$$F(y) = \frac{1}{2}\int_\Omega [y(T,x)^+]^2\, dx.$$

F n'est donc pas d'un type décrit dans la remarque 1. Montrons néanmoins qu'elle vérifie les hypothèses nécessaires. Il est clair que F est convexe, s.c.i. et a pour G-différentielle F'(y) définie par :

$$(19)\qquad \langle F'(y),\, z\rangle_{V'V} = \int_\Omega y(T,x)^+\, z(T,x)\, dx,$$

et on cherche $\alpha > 0$ tel que

$$(20)\quad \begin{cases} \frac{1}{2}\int_\Omega [y_2(T,x)^+]^2\, dx \geq \frac{1}{2}\int_\Omega [y_1(T,x)^+]^2\, dx + \int_\Omega y_1(T,x)^+(y_2(T,x) - y_1(T,x))dx + \\ + \alpha||F'(y_1) - F'(y_2)||^2_{V'}\ . \end{cases}$$

Majorons le dernier terme : de (19) on tire

$$|\langle F'(y_1) - F'(y_2), z\rangle|_{V'V} \leq \; ||y_1(T,.)^+ - y_2(T,.)^+||_{L^2(\Omega)} \times$$

$$\times \; ||z(T,.)||_{L^2(\Omega)},$$

et l'application $z : \to z(T,.)$ étant continue de V dans $L^2(\Omega)$ il existe $\gamma > 0$ tel que

$$|\langle F'(y_1) - F'(y_2), z\rangle|_{V'V} \leq \gamma \; ||y_1(T,.)^+ - y_2(T,.)^+||_{L^2(\Omega)} \; ||z||_V \; ;$$

ceci $\forall z \in V$, d'où

$$||F'(y_1) - F'(y_2)||_{V'} \leq \gamma ||y_1(T,.)^+ - y_2(T,.)^+||_{L^2(\Omega)} .$$

On va donc améliorer (20) en trouvant α_1 tel que

$$\frac{1}{2} \int_\Omega [y_2(T,x)^+]^2 \, dx \geq \frac{1}{2} \int_\Omega [y_1(T,x)^+]^2 \, dx +$$

$$+ \int_\Omega \; y_1(T,x)^+(y_2(T,x) - y_1(T,x))dx + \alpha_1 \int_\Omega [y_1(T,x)^+ - y_2(T,x)^+]^2 \, dx \; ;$$

il suffit de trouver α_1 tel que la fonction

$$\frac{1}{2} [y_2(T,x)^+]^2 - \frac{1}{2}[y_1(T,x)^+]^2 - y_1(T,x)^+ \; (y_2(T,x) - y_1(T,x)) -$$

$$- \alpha_1[y_1(T,x)^+ - y_2(T,x)^+]^2$$

soit presque partout positive. Il suffit de montrer que, si a, b $\in \mathbb{R}$ on a, pour un certain $\alpha_1 > 0$:

$$\frac{1}{2}(a^+)^2 - \frac{1}{2}(b^+)^2 - b^+(a - b) - \alpha_1(a^+ - b^+)^2 \geq 0.$$

Cette inégalité est vérifiée pour $\alpha_1 \leq 1/2$.

On a encore l'existence d'un contrôle optimal et des arguments analogues à ceux du cas distribués induisent la :

PROPOSITION 4. Si $c_2 > 0$ et c_1 est assez grand, on obtient les solutions du problème de contrôle (18) en minimisant le lagrangien augmenté

$$(21) \quad \begin{cases} S(y,u,p) = \frac{1}{2}\int_\Omega [y(T,x)^+]^2\, dx + \frac{N}{2}||u||^2_{L^2(\Sigma)} - \\ - \int_0^T \langle p, \frac{\partial y}{\partial t} + ay - f - Bu\rangle_{H^1(\Omega)H^1(\Omega)'} dt + c_1 \int_0^T ||\frac{\partial y}{\partial t} + ay - f - Bu||^2_{H^1(\Omega)'}\, dt + \\ + c_2 \int_0^T ||-\frac{\partial p}{\partial t} + ap||^2_{H^1(\Omega)'}\, dt + c_2||p(T,.) - y(T,.)^+||^2_{L^2(\Omega)}\, dt, \end{cases}$$

pour $(y,u,p) \in Z(0,T) \times L^2(\Sigma) \times Z(0,T)$ et sous les contraintes

$$\begin{cases} y(0,x) = 0 \text{ dans } \Omega, \\ u \in K. \end{cases}$$

II.5. Cas d'une équation d'état non affine

Si l'équation d'état n'est pas linéaire le problème n'est en général plus convexe. Dans ce cas les conditions d'optimalité obtenues sont nécessaires et non plus suffisantes. Notons d'ailleurs que le traitement des problèmes non convexes par des méthodes de lagrangien augmenté n'a été bien étudié qu'en dimension finie (voir R.T. Rockafellar [14]). Nous nous contenterons d'étudier la méthode sur l'exemple suivant. Soit $u \in W(0,T)$ et soit l'équation d'état :

$$(22) \quad \begin{cases} y(0,x) = y_o(x) \text{ dans } \Omega, \\ \frac{\partial}{\partial t} G(y) + ay = f + Bu\ ; \end{cases}$$

a et B étant définis comme précédemment, avec $y_o(x) \in L^2(\Omega)$, $f \in L^2(Q)$ et $u \in L^2(\Sigma)$; G est une fonction $R \to R$, de classe C^1, strictement croissante. Le problème de contrôle est

$$(23)\qquad \begin{cases} \text{minimiser } J(y,u) = \frac{1}{2}||y - y_d||^2_{L^2(Q)} + N||u||^2_{L^2(\Sigma)} + I_K(u), \\ \text{pour } (y,u) \in Z(0,T) \times L^2(\Sigma) \text{ et vérifiant (21).} \end{cases}$$

On peut résoudre ce type de problème par l'introduction d'un état adjoint (voir C. Saguez [15]). L'équation adjointe est :

$$\begin{cases} p(T,x) = 0 \text{ dans } \Omega, \\ - G'(y) \frac{\partial}{\partial t} p + ap = y - y_d, \end{cases}$$

où $G'(y)$ est la dérivée de G en y. Introduisons le lagrangien augmenté

$$(24)\qquad \begin{cases} S(y,u,p) = J(y,u) - \int_0^T \langle p, \frac{\partial}{\partial t} G(y) + ay - f - Bu\rangle_{H^1(\Omega)H^1(\Omega)'} dt + \\ + c_1 \int_0^T ||\frac{\partial}{\partial t} G(y) + ay - f - Bu||^2_{H^1(\Omega)'} dt + \\ + c_2 \int_0^T ||-G'(y)\frac{\partial p}{\partial t} + ap - (y - y_d)||^2_{H^1(\Omega)'} dt. \end{cases}$$

En l'absence de résultats théoriques, on examinera les résultats de la minimisation numérique de S sous les contraintes

$$\begin{cases} y(0,x) = y_o(x) \text{ dans } \Omega, \\ p(T,x) = 0. \end{cases}$$

Remarquons qu'il n'est plus nécessaire de résoudre l'équation d'état ce qui, dans le cas où elle est non linéaire, peut représenter un gain de temps de calcul important.

III. RESULTATS NUMERIQUES

III.1. Problème de contrôle et expression du lagrangien augmenté

On a étudié numériquement le problème (23) avec $T = 1$ et $\Omega =]0,1[$. Plus précisément l'équation d'état est :

$$(25)\quad \begin{cases} y(0,x) = y_o(x) \text{ pour } x \in]0,1[, \\ \dfrac{\partial}{\partial t} G(y) - \Delta y = f \text{ pour } (t,x) \in]0,1[\times]0,1[, \\ -\dfrac{\partial y}{\partial n}(t,0) = u(t), \\ -\dfrac{\partial y}{\partial n}(t,1) = 0, \end{cases}$$

avec $u(t) \in L^2(0,1)$; on a pris

$$\begin{cases} y_o(x) = 2,25 + 0,25\, x(x-2), \\ f(t,x) = -\,0,25\, x(x-2) - 1,25 - 2(1-t), \end{cases}$$

G étant précisé ultérieurement. Le problème de contrôle est :

$$(26)\quad \begin{cases} \text{minimiser} \\ J(y,u) = \int_0^1 \int_0^1 (y(t,x) - y_d(t,x))^2 \, dx\, dt + 2,5 \int_0^1 (u(t))^2 \, dt, \\ \text{pour } (y,u) \text{ vérifiant } (24). \end{cases}$$

Pour résoudre le problème (26) on va minimiser le lagrangien augmenté

$$\begin{aligned} S(y,u,p) = J(y,u) &- \int_0^1 \left\langle p, \frac{\partial G(y)}{\partial t} + ay - f - Bu \right\rangle_{H^1(\Omega) H^1(\Omega)'} dt + \\ &+ c_1 \int_0^1 \left\| \frac{\partial}{\partial t} G(y) + ay - f - Bu \right\|^2_{H^1(\Omega)'} + \\ &+ c_2 \int_0^1 \left\| - G'(y) \frac{\partial p}{\partial t} + ap - (y - y_d) \right\|^2_{H^1(\Omega)'} dt, \end{aligned}$$

Sous les contraintes

$$\begin{cases} y(0,x) = y_o(x) \quad x \in]0,1[, \\ p(1,x) = 0 \quad x \in]\,0,1[, \end{cases}$$

a et B étant les opérateurs linéaires de $H^1(\Omega)$ vers $H^1(\Omega)'$ et de $L^2(0,1)$ vers V' définis par

$$<a\psi,\phi>_{H^1(\Omega)'H^1(\Omega)} = \int_\Omega \nabla\phi\nabla\psi \, dx,$$

$$<Bu,\phi>_{H^1(\Omega)'H^1(\Omega)} = \int_0^1 u(t) \, \phi(t,0)dt.$$

Cas d'une équation d'état linéaire

On prend $G(y) = y$.

S est alors une fonctionnelle quadratique du problème de contrôle ; voir J.P. Yvon [16] . Numériquement, on se ramène à un problème sans contraintes en fixant certaines valeurs de y et p. On résoud donc, en fait, un système linéaire :

$$\text{grad } S = 0.$$

Néanmoins, vu la taille du système une résolution directe est difficile. Nous minimisons donc la fonctionnelle par un algorithme de gradient conjugué. En conséquence, le conditionnement de S jouera un rôle essentiel. C'est pourquoi, en raison de leur influence sur le conditionnement, on va brièvement préciser le mode de discrétisation et les normes choisies.

On choisit une semi-discrétisation en temps de l'équation d'état suivant un schéma implicite. Soit N le nombre de pas de temps et $\Delta t = \frac{1}{N}$, on cherche $y = (y_1, \ldots, y^N) \in [H^1(0,1)]^N$ solution de

$$\frac{y^{n+1}-y^n}{\Delta t} + ay^{n+1} = f^{n+1} + Bu^{n+1} \text{ pour } n = 0 \text{ à } N-1 \text{ ;}$$

y^o, f^n, u^n sont des approximations de $y_o(x)$, $f(t,x)$ et $u(t)$.

La discrétisation en espace s'effectue suivant une méthode de différences finies.

On prend comme norme pour y (idem pour p)

$$(27)\quad \begin{cases} \|y\|^2_{[H^1(0,1)]^N} = \sum_{i=1}^{N} \|y^i\|^2_{H^1(0,1)} \\ \text{et , } a > 0 \text{ étant donné :} \\ \|y^i\|^2_{H^1(0,1)} = \frac{1}{a\Delta t} \int_0^1 (y^i(x))^2\, dx + \int_0^1 (\frac{\partial y^i}{\partial x}(x))^2\, dx. \end{cases}$$

Remarquons que, après discrétisation en espace, y^i appartient à un espace de dimension finie et on peut prendre pour y^i la norme de L^2 qui est alors équivalente. Ce choix a conduit à de mauvais résultats numériques ; nous ne le détaillerons donc pas davantage. On note :

N	nombre de pas temps,
Nx	nombre de pas en espace,
NI	nombre d'itérations nécessaires pour satisfaire un test de convergence,
temps	le temps de calcul sur machine HB 68 de l'INRIA,
test	la valeur du test d'arrêt relatif sur la valeur de S,
a	paramètre intervenant dans la norme de y et p : voir (27).

On prendra toujours les conditions initiales :

$$\begin{cases} y = 0 \text{ sur }]0,1[\times]0,1[, \\ u = 0 \text{ sur }]0,1[, \\ p = 0 \text{ sur }]0,1[\times]0,1[. \end{cases}$$

<u>TABLEAU 1</u> Influence du choix de c_1 sur la vitesse de convergence. On prend $N = Nx = 10$, $a = 1$, test $= 10^{-3}$, $c_2 = 80$ et on donne NI fonction de c_1.

C_1	1	5	20	30	60	200
NI	ne converge pas	33	25	23	28	45

TABLEAU 2 Influence du choix de c_2 sur la vitesse de convergence. Mêmes valeurs que dans le Tableau 1 pour N, Nx, a, test ; on prend $c_1 = 30$.

c_2	5	10	80	200
NI	ne converge pas	29	23	27

L'algorithme est donc plus sensible au choix de c_1 qu'au choix de c_2.

TABLEAU 3 Influence du choix de N, nombre de pas de temps ; on prend Nx = 10, a = 1, test = 10^{-3}, $c_1 = 30$, $c_2 = 80$.

N	10	30	50
NI	23	54	77
temps	11,5 s	72 s	165 s

TABLEAU 4 Influence du choix de Nx. On prend N = 10, a = 1, test = 10^{-3}, $c_1 = 30$, $c_2 = 80$.

Nx	10	30	50
NI	23	28	29
temps	11,5 s	37 s	63 s

Les temps de calcul sont une fonction à peu près linéaire de Nx, ce qui est satisfaisant. Par contre ils augmentent très vite quant N croît.

TABLEAU 5 Influence du choix de la norme. On prend $N = Nx = 10$, test $= 10^{-3}$, $c_1 = 30$, $c_2 = 80$ et on fait varier a défini dans (27).

a	0,5	0,6	1	1,5	2
NI	29	24	22	24	30

Les essais effectués n'ont donc pas abouti à mettre en évidence une norme de $H^1(\Omega)$ mieux adaptée au problème (26) que la norme habituelle qui correspond au cas $a = 1$.

TABLEAU 6 Comparaison avec la méthode classique.

Rappelons que cette méthode consiste à considérer l'état comme fonction du contrôle. L'introduction de l'état adjoint permet d'obtenir le gradient du critère. On choisit ici de minimiser par une méthode de gradient conjugé. La méthode converge numériquement en trois itérations (après, le critère reste stationnaire). On donne une idéee des temps de calcul dans le Tableau 6 avec $Nx = 10$, N variant.

N	10	50	100	150
temps	0,6 s	3,3 s	6,6 s	10,4 s

Les temps de calcul sont proportionnels au nombre de pas de temps, ce qui n'est pas surprenant (puisque le nombre d'itérations est le même dans tous les cas, les calculs pour chaque itération étant proportionnels à N). La solution du problème de contrôle est obtenue ici beaucoup plus rapidement qu'en utilisant le lagrangien augmenté.

REMARQUE 5 On pourrait objecter que la convergence rapide de la méthode classique est dûe à l'importance du terme dépendant de u dans $J(y,u)$; en effet ce terme tend à améliorer le conditionnement de $J(y(u),u)$ (où $y(u)$ représente l'état associé au contrôle u). Toutefois, des essais numériques effectués en ne conservant dans J que le terme dépendant de l'état montrent encore une convergence assez rapide.

Cas d'une équation d'état non linéaire

On considère la fonction G définie par

$$\begin{cases} G(y) = y \text{ si } y \leq 1, \\ G(y) = \alpha(y-1) + 1 \text{ si } y \geq 1. \end{cases}$$

Pour $\alpha = 1$ on retrouve le cas précédent. Plus α s'éloigne de 1, plus l'équation d'état est non linéaire. On reprend les mêmes conditions initiales (toutes les inconnues égales à zéro). Le lagrangien augmenté est minimisé par une méthode de type Fletcher-Reeves. La solution trouvée n'est considérée comme correcte que si elle vérifie l'équation d'état et l'équation de l'état adjoint.

On prend $N = Nx = 10$, $a = 1$, test $= 2.10^{-3}$, $c_1 = 30$, $c_2 = 80$.

α	0,5	0,75	1	1,25	1,5	1,75
NI	divergence	28	22	34	43	divergence
temps		17 s	13 s	20 s	25 s	

La méthode donne satisfaction pour des petites non-linéarités seulement.

IV. CONCLUSION

La théorie de G. Di Pillo et L. Grippo a été étendue aux problèmes de contrôle de systèmes distribués avec critère convexe et équation d'état linéaire. Un exemple de contrôle de l'équation de la chaleur a été traité numériquement. La solution a été obtenue avec des temps de calcul beaucoup plus longs que ceux de la méthode classique. On peut voir là une confirmation de l'efficacité de la méthode du gradient réduit. Peut-être un choix judicieux de la norme permettra-t-il d'améliorer le conditionnement du lagrangien augmenté, et donc d'augmenter la vitesse de convergence. On peut aussi songer à une utilisation numérique de la fonctionnelle pénalisée exacte.

V. BIBLIOGRAPHIE

[1] D.P. BERTSEKAS
"Multipliers Methods : A Survey"
Automatica, Vol. 12, pp. 133-145. (1976).

[2] P.T. BOGGS AND J.W.TOLLE
"Augmented lagrangians wich are quadratic in the multiplier"
J.O.T.A., Vol. 31, n° 1, pp. 17-26. (1980).

[3] G. DI PILLO - L. GRIPPO
"The Multiplier Method for Optimal Control Problems of Parabolic Systems"
Appl. Math. Optim. 5, pp. 253-269. (1979).

[4] G. DI PILLO - L. GRIPPO
"A New Class of Augmented Lagrangians in Nonlinear Programming"
SIAM J. of Control and Optim., Vol. 17, n° 5, pp. 618-628. (1979).

[5] G. DI PILLO - L. GRIPPO - F. LAMPARIELLO
"A Computing technique for Solving Discrete-Time Optimal Control Problems"
2nd IFAC Workshop on Control Applications of Non Linear Programming and Optimisation. Munich (1980).

[6] I. EKELAND - R. TEMAM
"Analyse convexe et problèmes variationnels"
Dunod, Paris (1974).

[7] R. FLETCHER

"A class of methods for nonlinear programming with termination and convergence properties"
In Integer and Nonlinear Programming, ed. J. Abadie, North Holland, (1970).

[8] M. FORTIN - R. GLOWINSKI - F. THOMASSET

"Résolution numérique de problèmes aux limites par des méthodes de lagrangien augmenté "
Livre à paraître.

[9] D. GABAY

"Méthodes numériques pour l'optimisation non linéaire"
Thèse, Université Paris VI. (1979).

[10] R. GLOWINSKI

"Numerical methods for non-linear variational problems"
Tata Institute of fundamental Research, Bombay (1980).

[11] M.R. HESTENES

"Multiplier and gradient methods"
J.O.T.A. 4, pp. 303-320 (1969).

[12] J.L. LIONS

"Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles"
Dunod, Paris (1968).

[13] M.J.D. POWELL

"A method for nonlinear constraints in minimisation problems"
in Optimisation (R. Fletcher, ed), pp. 283-289. Academic Press, New York (1969).

[14] R.T. ROCKAFELLAR

"Augmented Lagrange Multiplier Functions and Duality in Nonconvex Programming"
SIAM J. of Control, Vol. 12, n° 2, pp. 268-285 (1974).

[15] C. SAGUEZ
"Contrôle optimal de systèmes à frontière libre"
Thèse, UTC Compiègne, (1980).

J.P. YVON
Application de la pénalisation à la résolution d'un problème de contrôle optimal"
Cahier de l'IRIA, n° 2 (1970).